ON THE CORRELATIONS, SELBERG INTEGRAL AND SYMMETRY OF SIEVE FUNCTIONS IN SHORT INTERVALS

by G.Coppola

Abstract. We study the arithmetic (real) function $f = g * \mathbf{1}$, with g "essentially bounded" and supported over the integers of $[1, Q]$. In particular, we obtain non-trivial bounds, through f "correlations", for the "Selberg integral" and the "symmetry integral" of f in almost all short intervals $[x - h, x + h]$, $N \le x \le 2N$, beyond the "classical" level, up to level of distribution, say, $\lambda = \log Q / \log N < 2/3$ (for enough large h). This time we don't apply Large Sieve inequality, as in our paper [C-S]. Precisely, our method is completely elementary.

1. Introduction and statement of the results.

We study "SIEVE FUNCTIONS" , i.e. real arithmetic functions $f = g*\mathbf{1}$ (see hypotheses on g in the sequel), in almost all the short intervals $[x-h, x+h]$ (i.e., almost all stands $\forall x \in [N, 2N]$, except $o(N)$ of them and short means, say, $h \to \infty$ and $h = o(N)$, as $N \to \infty$). Here, as usual, $\mathbf{1}(n) = 1$ is the constant-1 arithmetic function and $*$ is the Dirichlet product (esp., [T]). In order to study the sum of f values in a.a. (abbreviates almost all, now on) the intervals $[x - h, x + h]$, we define (in analogy with the classical Selberg integral, see [C-S]) the "SELBERG INTEGRAL" of f as: $J_f(N,h) \overset{def}{=} \int_N^{2N} \Big| \sum_{0<|n-x|\le h} f(n) - M_f(2h)\Big|^2 dx$, where (from heuristics in accordance with the classical case) we expect the "mean-value" to be $M_f(2h) \overset{def}{=} 2h \sum_d g(d)/d$ (that converges in interesting cases and under our hypotheses on g, see the sequel; also, $d \le 2N + h$, here). Furthermore, this definition comes from what the "natural" choice of $M_f(2h)$ is (recall $[\ ] =$ INTEGER PART):

$$2h\Big(\frac{1}{x}\sum_{n\le x} f(n)\Big) = \frac{2h}{x}\sum_d g(d)\left[\frac{x}{d}\right] = 2h\sum_d \frac{g(d)}{d} + \mathcal{O}\Big(\frac{h}{x}\sum_{d\le Q}|g(d)|\Big),$$

in fact, when $f = g * \mathbf{1}$, $g(q) = 0$ for $q > Q$. Assuming Q smaller than x (in the sequel), we recover $M_f(2h)$. Selberg integral counts the values of f in a.a. $[x-h, x+h]$. We study their symmetry through the "SYMMETRY INTEGRAL" of f (here $\mathrm{sgn}(0) \overset{def}{=} 0$, $\mathrm{sgn}(r) \overset{def}{=} \frac{|r|}{r}$, $\forall r \ne 0$): $I_f(N,h) \overset{def}{=} \int_N^{2N} \Big| \sum_{|n-x|\le h} \mathrm{sgn}(n-x) f(n)\Big|^2 dx.$

We'll generalize the results given in [C-S] for these integrals, applying the Large Sieve inequality, in the case $g = \mathbf{1}$ of the divisor function $d = \mathbf{1} * \mathbf{1}$. We point out that the procedure given there works, as well, for more general g to bound I_f; but fails in the case of J_f, whenever g is not constant (i.e., the Dirichlet "flipping" of the divisors can't be applied). Here, we give another approach valid for both integrals, even for non-constant g. It is based on the "correlations" of f. The CORRELATION OF f IS defined as ($\forall a \in \mathbb{Z}, a \ne 0$)

$$\mathcal{C}_f(a) \overset{def}{=} \sum_{n\sim N} f(n)f(n-a) = \sum_{\ell|a}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\left(\left[\frac{2N}{\ell d}\right] - \left[\frac{N}{\ell d}\right]\right) + R_f(a)$$

(hereon $x \sim X$ is $X < x \le 2X$), where, through the orthogonality of additive characters [V] as in Lemma 3 (as usual, we will always write $e(\theta) \overset{def}{=} e^{2\pi i\theta}$, $\forall \theta \in \mathbb{R}$ and $e_q(m) \overset{def}{=} e(m/q)$, $\forall q \in \mathbb{N}$, $\forall m \in \mathbb{Z}$), say,

$$R_f(a) \overset{def}{=} \sum_{\ell|a}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\ne 0} e_q(-ja/\ell) \sum_{m\sim \frac{N}{\ell d}} e_q(jdm)$$

(here, and in the following, $j \ne 0$ means that j describes exactly once all classes (mod q), except $j \equiv 0(q)$); and I_f is a sum (see Lemma 1) of these correlations, weighted with W (name from the shape), W EVEN,

$$W(a) \overset{def}{=} \begin{cases} 2h - 3a & \text{if } 0 \le a \le h \\ a - 2h & \text{if } h \le a \le 2h \\ 0 & \text{if } a > 2h \end{cases} \qquad \Longrightarrow \qquad \sum_{a\in\mathbb{Z}} W(a) = 0.$$

In complete analogy, Lemma 2 gives the Selberg integral $J_f(N,h)$ as a weighted sum of correlations, with (Selberg) weight $S(a) \overset{def}{=} \max(2h - |a|, 0)$. Notice that S is always non-negative (while W oscillates in sign).

Mathematics Subject Classification (2000) : $11N37, 11N25$.

(Here, as usual, $F=o(G) \overset{def}{\iff} \lim F/G=0$ and $F=\mathcal{O}(G) \overset{def}{\iff} \exists c>0 : |F|\le c\,G$ are Landau's notation. Also, when c depends on ε, we'll write $F=\mathcal{O}_\varepsilon(G)$ or, like Vinogradov, $F\ll_\varepsilon G$). We call an arithmetical function ESSENTIALLY BOUNDED when, $\forall \varepsilon>0$, its $n-$th value is at most $\mathcal{O}_\varepsilon(n^\varepsilon)$ and we'll write $\lll_\varepsilon 1$; i.e.,

$$F(N) \lll_\varepsilon G(N) \overset{def}{\iff} \forall \varepsilon>0 \;\; F(N)\ll_\varepsilon N^\varepsilon G(N) \quad (\text{as } N\to\infty)$$

e.g., the divisor function $d(n)$ is essentially bounded (like many other number-theoretic f) and we remark that $f=g*\mathbf{1}$ is essentially bounded if and only if g is (from Möbius inversion, see [D]). From Lemma 2, applying Lemma 3 to f correlations, together with $\sum_a S(a\ell)=4h^2/\ell+\mathcal{O}(h)$, uniformly $\forall \ell\in\mathbb{N}$ (like in (1), see Lemma 4 proof), we get

$$J_f(N,h)=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\sum_a S(a\ell)R_f(a)+\mathcal{O}_\varepsilon\left(N^\varepsilon\left(Nh+Qh^2\right)\right)$$

In fact, (compare the discussion about $M_f(2h)$, above)

$$f\lll_\varepsilon 1 \Rightarrow M_f(2h)\lll_\varepsilon h, \;\; \sum_{n\sim N} f(n)=\sum_d g(d)\left(\left[\frac{2N}{d}\right]-\left[\frac{N}{d}\right]\right)=N\sum_d \frac{g(d)}{d}+\mathcal{O}_\varepsilon\left(N^\varepsilon Q\right).$$

We recall $\|r\| \overset{def}{=} \min_{n\in\mathbb{Z}}|r-n|$ is the DISTANCE FROM INTEGERS. We abbreviate $n\equiv a(\bmod\ q)$ with $n\equiv a(q)$.

We give our main result.

THEOREM. *Let* $N,h,Q\in\mathbb{N}$, *be such that* $h\to\infty$, $Q\ll N$ *and* $h=o(N)$, *as* $N\to\infty$. *Let* $f:\mathbb{N}\to\mathbb{R}$ *be essentially bounded, with* $f=g*\mathbf{1}$ *and* $g(q)=0\ \forall q>Q$. *Then*

$$J_f(N,h)\lll_\varepsilon Nh+h^3+Q^2h+Qh^2; \qquad I_f(N,h)\lll_\varepsilon Nh+h^3+Q^2h+Qh^2.$$

Also, only for the symmetry integral $I_f(N,h)$,

$$I_f(N,h)=2\sum_a S(a)\left(\mathcal{C}_f(a)-\mathcal{C}_f(a+h)\right)+\mathcal{O}_\varepsilon\left(N^\varepsilon(Nh+h^3)\right).$$

Remark. We explicitly point out that our Theorem implies non-trivial estimates $J_f(N,h)\ll \frac{Nh^2}{N^\varepsilon}$ and $I_f(N,h)\ll \frac{Nh^2}{N^\varepsilon}$ for both integrals, with LEVEL OF DISTRIBUTION, say, $\frac{\log Q}{\log N}\overset{def}{=}\lambda<\frac{1+\theta}{2}$, where, say, $\theta\overset{def}{=}(\log h)/(\log N)$ is the WIDTH; hence, level up to 2/3, when the width is above 1/3. (The same result can also be achieved with the method of [C-S], but only for I_f.)

In fact, an immediate consequence of our Theorem is the following

COROLLARY. *Let* $0<\theta<1$, $0\le\lambda<\frac{1+\theta}{2}$ *and* $N,h,Q\in\mathbb{N}$, *be such that* $N^\theta\ll h\ll N^\theta$, $N^\lambda\ll Q\ll N^\lambda$, *as* $N\to\infty$. *Let* $f:\mathbb{N}\to\mathbb{R}$ *be essentially bounded, with* $f=g*\mathbf{1}$ *and* $g(q)=0\ \forall q>Q$. *Then* $\exists\varepsilon_0=\varepsilon_0(\theta,\lambda)>0$ (*depending only on* θ,λ) *such that*

$$J_f(N,h)\ll_{\varepsilon_0} Nh^2N^{-\varepsilon_0}, \qquad I_f(N,h)\ll_{\varepsilon_0} Nh^2N^{-\varepsilon_0}.$$

The paper is organized as follows:

◇ we will give our Lemmas in the next section;

◇ then we will prove our Theorem in section 3.

2. Lemmas.

Lemma 1. *Let* $N,h\in\mathbb{N}$, *with* $h\to\infty$ *and* $h=o(N)$ *as* $N\to\infty$. *If* $f:\mathbb{N}\to\mathbb{R}$ *has* $\|f\|_\infty=\max_{n\le 2N+h}|f(n)|$,

$$\int_N^{2N}\Big|\sum_{|n-x|\le h}\mathrm{sgn}(n-x)f(n)\Big|^2dx=\sum_a W(a)\mathcal{C}_f(a)+\mathcal{O}\left(h^3\|f\|_\infty^2\right).$$

Proof. This is a kind of dispersion method, without "expected mean":the main term "vanishes". Use f REAL:

$$I_f(N,h)=D_f(N,h)+2\sum\sum_{N-h<n_1<n_2\le 2N+h}f(n_1)f(n_2)\int_{x\sim N,|x-n_1|\le h,|x-n_2|\le h}\mathrm{sgn}(x-n_1)\mathrm{sgn}(x-n_2)dx;$$

here (I_f is the integral above and) $D_f(N,h):=\sum_{N-h<n\le 2N+h}f^2(n)\int_{N<x\le 2N,0<|x-n|\le h}dx=$

$$=\sum_{N+h<n\le 2N-h}f^2(n)\int_{|x-n|\le h}dx+\mathcal{O}\left(h\|f\|_\infty^2\left(\sum_{|n-N|\le h}1+\sum_{|n-2N|\le h}1\right)\right)=W(0)\mathcal{C}_f(0)+\mathcal{O}\left(h^2\|f\|_\infty^2\right)$$

is the DIAGONAL. The remainder, here, is (negligible) in the second one. Since (for $a>0$)

$$\mathcal{C}_f(-a)=\sum_{n\sim N}f(n)f(n+a)=\sum_{N+a<m\le 2N+a}f(m-a)f(m)=\mathcal{C}_f(a)+\mathcal{O}\left(a\|f\|_\infty^2\right),$$

W EVEN and $W(a)\ll h\ \Rightarrow\ \sum_{0<a\le 2h}W(a)\mathcal{C}_f(-a)=\sum_{0<a\le 2h}W(a)\mathcal{C}_f(a)+\mathcal{O}\left(h^3\|f\|_\infty^2\right)$, we confine to:

$$(*)\qquad I_f(N,h)-D_f(N,h):=W(0)\mathcal{C}_f(0)+2\sum_{0<a\le 2h}W(a)\mathcal{C}_f(a)+E_f(N,h),\ \text{say},\ E_f(N,h)\ll h^3\|f\|_\infty^2.$$

The left-hand side, changing variables, namely $n=n_1$, $a=n_2-n_1$, $s=x-n_1$, is (introducing the remainders which shall take part of the final $E_f(N,h)$, here)

$$2\sum_{N-h<n<2N+h}f(n)\sum_{0<a\le 2h,a\le 2N+h-n}f(n+a)\int_{N-n<s\le 2N-n,|s|\le h,|s-a|\le h}\mathrm{sgn}(s)\mathrm{sgn}(s-a)ds=$$

$$=2\sum_{N-h<n\le 2N-h}f(n)\sum_{0<a\le 2h}f(n+a)\int_{s>N-n,|s|\le h,|s-a|\le h}\mathrm{sgn}(s)\mathrm{sgn}(s-a)ds+E_1=$$

$$=2\sum_{N+h<n\le 2N-h}f(n)\sum_{0<a\le 2h}f(n+a)\int_{|s|\le h,|s-a|\le h}\mathrm{sgn}(s)\mathrm{sgn}(s-a)ds+E_1+E_2=$$

$$=2\sum_{N<n\le 2N}f(n)\sum_{0<a\le 2h}f(n+a)W(a)+E_1+E_2+E_3,\ \ W(a):=\int_{\substack{|s|\le h\\|s-a|\le h}}\mathrm{sgn}(s)\mathrm{sgn}(s-a)ds\ll h,$$

whence $E_3\ll\left(\sum_{N<n\le N+h}+\sum_{2N-h<n\le 2N}\right)|f(n)|\sum_{0<a\le 2h}|f(n+a)|h\ll h^3\|f\|_\infty^2$ is a "TAIL", like:

$$E_1\ll\sum_{|n-2N|\le h}|f(n)|\sum_{0<a\le 2h}|f(n+a)|h,\ E_2\ll\sum_{|n-N|\le h}|f(n)|\sum_{0<a\le 2h}|f(n+a)|h\ \text{ are }\ll h^3\|f\|_\infty^2.\ \square$$

Lemma 2. *Let* $N, h \in \mathbb{N}$, *with* $h \to \infty$ *and* $h = o(N)$ *as* $N \to \infty$. *If* $f : \mathbb{N} \to \mathbb{R}$ *has* $\|f\|_\infty = \max_{n \le 2N+h} |f(n)|$,

$$\int_N^{2N} \Big| \sum_{0<|n-x|\le h} f(n) - M_f(2h) \Big|^2 dx = \sum_a S(a)\mathfrak{C}_f(a) - 4M_f(2h)h \sum_{n\sim N} f(n) + M_f^2(2h)N +$$

$$+\mathcal{O}\left(h^3 \, \|f\|_\infty^2 + h^2 \|f\|_\infty |M_f(2h)|\right).$$

Proof. This is a direct application of dispersion method [L]. Use f REAL (ignoring, now, sets of measure zero):

$$J_f(N,h) = D_f(N,h) + 2 \sum\sum_{N-h<n_1<n_2\le 2N+h} f(n_1)f(n_2) \int_{x\sim N, |x-n_1|\le h, |x-n_2|\le h} dx -$$

$$-2M_f(2h) \sum_{N-h<n\le 2N+h} f(n) \int_{x\sim N, |x-n|\le h} dx + M_f^2(2h) \int_N^{2N} dx =$$

$$= D_f(N,h) + 2 \sum\sum_{N-h<n_1<n_2\le 2N+h} f(n_1)f(n_2) \int_{x\sim N, |x-n_1|\le h, |x-n_2|\le h} dx - 4hM_f(2h) \sum_{n\sim N} f(n) + M_f^2(2h)N,$$

save an error which is $\mathcal{O}(|M_f(2h)|h^2\|f\|_\infty)$; here ($J_f$ is the integral above and)

$$D_f(N,h) = \sum_{N-h<n\le 2N+h} f^2(n) \int_{x\sim N, 0<|x-n|\le h} dx =$$

$$= \sum_{N<n\le 2N} f^2(n) \int_{0<|x-n|\le h} dx + \mathcal{O}\left(h\|f\|_\infty^2 \left(\sum_{|n-N|\le h} 1 + \sum_{|n-2N|\le h} 1 \right) \right) = S(0)\mathfrak{C}_f(0) + \mathcal{O}\left(h^2 \|f\|_\infty^2\right)$$

is the same diagonal (with same negligible remainder) of Lemma 1. In fact, we closely follow its proof; due to: S EVEN and $S(a) \ll h \Rightarrow \sum_{0<a\le 2h} S(a)\mathfrak{C}_f(-a) = \sum_{0<a\le 2h} S(a)\mathfrak{C}_f(a) + \mathcal{O}\left(h^3 \|f\|_\infty^2\right)$, we confine to

$$(*) \qquad \sum\sum_{N-h<n_1<n_2\le 2N+h} f(n_1)f(n_2) \int_{x\sim N, |x-n_1|\le h, |x-n_2|\le h} dx - \sum_{0<a\le 2h} S(a)\mathfrak{C}_f(a) := E_f(N,h) \overset{\text{say}}{\ll} h^3 \|f\|_\infty^2.$$

The left-hand side, changing variables, namely $n = n_1$, $a = n_2 - n_1$, $s = x - n_1$, is (see Lemma 1 proof)

$$\sum_{N-h<n<2N+h} f(n) \sum_{0<a\le 2h, a\le 2N+h-n} f(n+a) \int_{N-n<s\le 2N-n, |s|\le h, |s-a|\le h} ds =$$

$$= \sum_{N<n\le 2N} f(n) \sum_{0<a\le 2h} f(n+a)S(a) + E_1 + E_2 + E_3, \quad S(a) := \int_{\substack{|s|\le h \\ |s-a|\le h}} ds \ll h,$$

whence $E_3 \ll \left(\sum_{N<n\le N+h} + \sum_{2N-h<n\le 2N} \right) |f(n)| \sum_{0<a\le 2h} |f(n+a)|h \ll h^3 \|f\|_\infty^2$ is a "TAIL", like:

$$E_1 \ll \sum_{|n-2N|\le h} |f(n)| \sum_{0<a\le 2h} |f(n+a)|h, \quad E_2 \ll \sum_{|n-N|\le h} |f(n)| \sum_{0<a\le 2h} |f(n+a)|h \text{ are } \ll h^3 \|f\|_\infty^2. \ \square$$

Lemma 3. *Let* $N,h,Q\in\mathbb{N}$, *where* $h\to\infty$, $h=o(N)$ *and* $Q\ll N$, *as* $N\to\infty$. *Let* $f=g*\mathbf{1}$, *where* $g:\mathbb{N}\to\mathbb{R}$, *with* $q>Q \Rightarrow g(q)=0$. *Then*

$$a\neq 0\Rightarrow \mathcal{C}_f(a)=\sum_{\ell|a}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\left(\left[\frac{2N}{\ell d}\right]-\left[\frac{N}{\ell d}\right]\right)+R_f(a),\ \textit{where, say, as in the introduction}$$

$$R_f(a)=\sum_{\ell|a}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\neq 0}e_q(-ja/\ell)\sum_{m\sim\frac{N}{\ell d}}e_q(jdm),\qquad \forall a\neq 0.$$

Also, every weight function $K:\mathbb{N}\to\mathbb{C}$, K EVEN, *with* $K(0)=2h$, *gives*

$$\sum_a K(a)R_f(a)=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\neq 0}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}\sum_{a\neq 0}K(a\ell)e_q(ja)+2h\mathcal{C}_f(0).$$

Proof. We'll always assume a non-zero. First of all, we start from the correlation, that is:

$$\mathcal{C}_f(a)=\sum_{n\sim N}f(n)f(n-a)=\sum_d\sum_q g(d)g(q)\sum_{\substack{N<n\le 2N\\ n\equiv 0(d)\\ n\equiv a(q)}}1=\sum_{(d,q)|a}\sum g(d)g(q)\sum_{\substack{\frac{N}{d}<m\le\frac{2N}{d}\\ md\equiv a(q)}}1,$$

since last congruence is solveable if and only if the GCD (d,q) divides a; changing variables, this is

$$\sum_{\ell|a}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\sum_{\substack{\frac{N}{\ell d}<m\le\frac{2N}{\ell d}\\ md\equiv\frac{a}{\ell}(q)}}1=\sum_{\ell|a}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\left(\left[\frac{2N}{\ell d}\right]-\left[\frac{N}{\ell d}\right]\right)+R_f(a),$$

using the orthogonality of additive characters (see [V]): here $R_f(a)$ is as above; summing on a with K,

$$\sum_{a\neq 0}K(a)R_f(a)=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\neq 0}\sum_{m\sim\frac{N}{\ell d}}e_q(jdm)\sum_{b\neq 0}K(b\ell)e_q(jb)=$$

(using K even, here)

$$=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\neq 0}\sum_{m\sim\frac{N}{\ell d}}e_q(jdm)\sum_{a\neq 0}K(a\ell)\cos\frac{2\pi ja}{q}=$$

$$=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\neq 0}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}\sum_{a\neq 0}K(a\ell)\cos\frac{2\pi ja}{q}=$$

$$=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\neq 0}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}\sum_{a\neq 0}K(a\ell)e_q(ja).$$

(We used once more K even, here.) Then, the thesis, adding the term $K(0)\mathcal{C}_f(0)=2h\mathcal{C}_f(0)$. ◻

Remark. We explicitly point out that, in our hypotheses on f (i.e., real and essentially bounded)

$$2h\mathcal{C}_f(0)=2h\sum_{n\sim N}f^2(n)\lll_\varepsilon Nh,$$

a trivial estimate which will be useful in future occurrences.

Lemma 4. *Defining,* $\forall h \in \mathbb{N}$, *the weight* W *as above, we have,* $\forall q \in \mathbb{N}$, $\forall \beta \notin \mathbb{Z}$, $\forall \ell \in \mathbb{N}$, $\forall \alpha \in \mathbb{R}$,

$$(1) \qquad \sum_{\substack{a \\ a\equiv 0 (\bmod q)}} W(a) = 2q \left\| \frac{h}{q} \right\|; \quad \sum_{0\le |a| \le 2h} W(a)e(a\beta) = \frac{4\sin^4(\pi h\beta)}{\sin^2(\pi\beta)}; \quad \sum_b W(\ell b)e(b\alpha) \ge 0.$$

Also, more in general (in the same hypotheses), abbreviating $E_X(\beta) \overset{def}{=} \sum_{0\le|a|\le X} e(a\beta)$, *we have*

$$(2) \qquad \frac{1}{\ell}\sum_a W(a\ell)e(a\beta) = \frac{4\sin^2 \pi\beta[\frac{h}{\ell}] - \sin^2 \pi\beta[\frac{2h}{\ell}]}{\sin^2 \pi\beta} + 4\left\{\frac{h}{\ell}\right\} E_{\frac{h}{\ell}}(\beta) - \left\{\frac{2h}{\ell}\right\} E_{\frac{2h}{\ell}}(\beta) =$$

$$= 2\left(1-\cos\left(2\pi\beta\left[\frac{h}{\ell}\right]\right)\right) \sum_{0\le|a|\le\frac{h}{\ell}} \left(\left[\frac{h}{\ell}\right] - |a|\right)e(a\beta) - \left(2\left\{\frac{h}{\ell}\right\} - \left\{\frac{2h}{\ell}\right\}\right) E_{2[\frac{h}{\ell}]}(\beta) +$$

$$+4\left\{\frac{h}{\ell}\right\} E_{\frac{h}{\ell}}(\beta) - \left\{\frac{2h}{\ell}\right\} E_{\frac{2h}{\ell}}(\beta).$$

Proof. Hereon $n \le X$ in a sum means $1 \le n \le X$. We will prove (1), even if it's a special case of (2);

$$\sum_{\substack{a \\ a\equiv 0 (\bmod q)}} W(a) = 2h + 4h\left(\left[\frac{h}{q}\right] - \left[\frac{2h}{q}\right] + \left[\frac{h}{q}\right]\right) + 2q\left(\sum_{\frac{h}{q}<b\le\frac{2h}{q}} b - 3\sum_{b\le\frac{h}{q}} b\right)$$

$$= q\left(\left\{\frac{2h}{q}\right\}^2 - 4\left\{\frac{h}{q}\right\}^2 - \left\{\frac{2h}{q}\right\} + 4\left\{\frac{h}{q}\right\}\right).$$

Using $\forall\alpha \in \mathbb{R}$ that $\{2\alpha\} = \{2\{\alpha\}\} = \begin{cases} 2\{\alpha\} & \text{if } 0 \le \{\alpha\} < 1/2 \\ 2\{\alpha\} - 1 & \text{if } 1/2 \le \{\alpha\} < 1 \end{cases}$ we get the first.

We come, now, to the second: $\sum_{0\le|a|\le 2h} W(a)e(a\beta) = 2h + 2\sum_{a\le 2h} W(a)\cos 2\pi a\beta = 2h + 2\Sigma$, say; then, partial summation gives $\Sigma = 4\sum_{a\le h} C_a(\beta) - 4C_h(\beta) - \sum_{a\le 2h} C_a(\beta) + C_{2h}(\beta)$, say, where $\forall X \in \mathbb{N}, \forall\theta \notin \mathbb{Z}$

$$C_X(\theta) \overset{def}{=} \sum_{n\le X}\cos(2\pi n\theta) = \frac{\sin(2\pi\theta X)}{2\tan(\pi\theta)} - \frac{1-\cos(2\pi\theta X)}{2} \quad \text{(a well-known formula)}$$

to get $\Sigma = 2\cot(\pi\beta)\sum_{a\le h}\sin(2\pi a\beta) - 2h - 2C_h(\beta) - \frac{1}{2}\cot(\pi\beta)\sum_{a\le 2h}\sin(2\pi a\beta) + \frac{1}{2}C_{2h}(\beta) + h =$

$$= \cot(\pi\beta)\left(2S_h(\beta) - \frac{1}{2}S_{2h}(\beta)\right) - h - 2C_h(\beta) + \frac{1}{2}C_{2h}(\beta), \text{ say, } \forall X \in \mathbb{N}, \forall\theta \notin \mathbb{Z}$$

$$S_X(\theta) \overset{def}{=} \sum_{n\le X}\sin(2\pi n\theta) = \frac{\sin^2(\pi\theta X)}{\tan(\pi\theta)} + \frac{\sin(2\pi\theta X)}{2}$$

Then, since $\frac{1-\cos(2\pi\beta X)}{2} = \sin^2(\pi\beta X)$, both for $X = h$ and $X = 2h$,

$$\Sigma = \cot^2(\pi\beta)\left(2\sin^2(\pi\beta h) - \frac{1}{2}\sin^2(2\pi\beta h)\right) + 2\sin^2(\pi\beta h) - \frac{1}{2}\sin^2(2\pi\beta h) - h =$$

$$= 2\cot^2(\pi\beta)\left(1-\cos^2(\pi h\beta)\right)\sin^2(\pi h\beta) + 2\sin^2(\pi h\beta)\left(1-\cos^2(\pi h\beta)\right) - h =$$

$$= 2\left(\cot^2(\pi\beta)+1\right)\sin^4(\pi h\beta) - h = \frac{2\sin^4(\pi h\beta)}{\sin^2(\pi\beta)} - h.$$

This gives the second. Finally, the third follows from: $\forall \ell \in \mathbb{N}$

$$\sum_b W(\ell b)e(b\alpha) \ge 0 \ \ \forall \alpha \in \mathbb{R} \Longleftrightarrow \sum_{a\equiv 0 (\mathrm{mod}\ \ell)} W(a)e(a\beta) \ge 0 \ \ \forall \beta \in \mathbb{R}$$

which, using the orthogonality of additive characters [V] and $\sum_a W(a)e(a\beta) \ge 0 \ \ \forall \beta \in \mathbb{R}$, is

$$\sum_{a\equiv 0 (\mathrm{mod}\ \ell)} W(a)e(a\beta) = \frac{1}{\ell}\sum_{j\le \ell}\sum_a W(a)e(a\beta)e_\ell(ja) = \frac{1}{\ell}\sum_{j\le \ell}\sum_a W(a)e\left(a\left(\beta+\frac{j}{\ell}\right)\right) \ge 0.$$

(We explicitly remark that this last property isn't "visible" from (2): not an immediate consequence.)

We come, now, to (2): $\displaystyle\sum_{0\le |a|\le 2h} W(a\ell)e(a\beta) = 2h + 2\sum_{a\le \frac{2h}{\ell}} W(a\ell)\cos 2\pi a\beta = 2h + 2\Sigma_\ell$, say; then,

$$\frac{1}{\ell}\Sigma_\ell = 4\sum_{a\le\left[\frac{h}{\ell}\right]} C_a(\beta) - 4C_{\left[\frac{h}{\ell}\right]}(\beta) - \sum_{a\le\left[\frac{2h}{\ell}\right]} C_a(\beta) + C_{\left[\frac{2h}{\ell}\right]}(\beta) + \left(4\left\{\frac{h}{\ell}\right\}C_{\left[\frac{h}{\ell}\right]}(\beta) - \left\{\frac{2h}{\ell}\right\}C_{\left[\frac{2h}{\ell}\right]}(\beta)\right),$$

from partial summation (the term in brackets isn't present whenever $\ell = 1$); then, (see above formulas)

$$\frac{1}{\ell}\Sigma_\ell = \cot(\pi\beta)\left(2S_{\left[\frac{h}{\ell}\right]}(\beta) - \frac{1}{2}S_{\left[\frac{2h}{\ell}\right]}(\beta)\right) - \left(2\left[\frac{h}{\ell}\right] - \frac{1}{2}\left[\frac{2h}{\ell}\right]\right) - 2C_{\left[\frac{h}{\ell}\right]}(\beta) + \frac{1}{2}C_{\left[\frac{2h}{\ell}\right]}(\beta) +$$
$$+\left(4\left\{\frac{h}{\ell}\right\}C_{\left[\frac{h}{\ell}\right]}(\beta) - \left\{\frac{2h}{\ell}\right\}C_{\left[\frac{2h}{\ell}\right]}(\beta)\right),$$

i.e.

$$\Sigma_\ell = \frac{2\sin^2\pi\beta\left[\frac{h}{\ell}\right] - \frac{1}{2}\sin^2\pi\beta\left[\frac{2h}{\ell}\right]}{\sin^2(\pi\beta)}\ell - h + \left(2\left\{\frac{h}{\ell}\right\}\left(1+2C_{\left[\frac{h}{\ell}\right]}(\beta)\right) - \frac{1}{2}\left\{\frac{2h}{\ell}\right\}\left(1+2C_{\left[\frac{2h}{\ell}\right]}(\beta)\right)\right)\ell;$$

hence,

$$\frac{1}{\ell}\sum_a W(a\ell)e(a\beta) = \frac{4\sin^2\pi\beta\left[\frac{h}{\ell}\right] - \sin^2\pi\beta\left[\frac{2h}{\ell}\right]}{\sin^2(\pi\beta)} + \left(4\left\{\frac{h}{\ell}\right\}\sum_{0\le|a|\le\frac{h}{\ell}} e(a\beta) - \left\{\frac{2h}{\ell}\right\}\sum_{0\le|a|\le\frac{2h}{\ell}} e(a\beta)\right);$$

we distinguish two cases: first, $0 \le \left\{\frac{h}{\ell}\right\} < \frac{1}{2}$ and, then, $\frac{1}{2} \le \left\{\frac{h}{\ell}\right\} < 1$. In the first, we have $\left[\frac{2h}{\ell}\right] = 2\left[\frac{h}{\ell}\right]$:

$$\frac{1}{\ell}\sum_a W(a\ell)e(a\beta) = \frac{4\sin^4\pi\beta\left[\frac{h}{\ell}\right]}{\sin^2(\pi\beta)} + \left(4\left\{\frac{h}{\ell}\right\}\sum_{0\le|a|\le\frac{h}{\ell}} e(a\beta) - \left\{\frac{2h}{\ell}\right\}\sum_{0\le|a|\le\frac{2h}{\ell}} e(a\beta)\right),$$

while in the second case we have $\left[\frac{2h}{\ell}\right] = 2\left[\frac{h}{\ell}\right] + 1$, so join (only for $2\left\{\frac{h}{\ell}\right\} - \left\{\frac{2h}{\ell}\right\} = 1$) the term

$$-\cos 4\pi\beta\left[\frac{h}{\ell}\right] - \sin 4\pi\beta\left[\frac{h}{\ell}\right]\cot(\pi\beta) = \left(\text{use the formula for } C_{2\left[\frac{h}{\ell}\right]}(\beta), \text{ here}\right)$$

$$= -2\left(\sum_{a\le 2\left[\frac{h}{\ell}\right]}\cos(2\pi a\beta) + \frac{1}{2}\right) = -\sum_{0\le|a|\le 2\left[\frac{h}{\ell}\right]} e(a\beta) = -\left(2\left\{\frac{h}{\ell}\right\} - \left\{\frac{2h}{\ell}\right\}\right)\sum_{0\le|a|\le 2\left[\frac{h}{\ell}\right]} e(a\beta). \ \square$$

3. Proof of the Theorem.
We will ignore the $R_g(N,h)$ that are $\lll_\varepsilon Nh+h^3$(Good remainders!). Linking the Lemmas,

$$I_f(N,h)=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\ne 0}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}\sum_{a\ne 0}W(a\ell)e_q(ja)$$

(save $\lll_\varepsilon R_g(N,h)$, hereon); and using Lemma 2 instead of Lemma 1, see the introduction,

$$J_f(N,h)=\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\ne 0}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}\sum_{a\ne 0}S(a\ell)e_q(ja)$$

Then, we'll show, for each K like in Lemma 3, supported in $[-2h,2h]$, where uniformly bounded as $K\ll h$,

$$(*)\quad T_g(N,h)\stackrel{def}{=}\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j\ne 0}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}\sum_{a}K(a\ell)e_q(ja)\lll_\varepsilon Qh^2+Q^2h+R_g(N,h)$$

In fact, we reintroduce terms with $a=0$ (here $K(0)=2h$), with contributes ($\sum^*$ =coprime to d)

$$2h\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)\frac{g(\ell q)}{q}\sum_{j\ne 0}\sum_{m\sim\frac{N}{\ell d}}e_q(jdm)=2h\sum_{\ell\le 2h}\sum_{d}g(\ell d)\sum_{m\sim\frac{N}{\ell d}}\left(\sum_{q|dm}{}^{*}g(\ell q)-\sum_{q}{}^{*}\frac{g(\ell q)}{q}\right)\lll_\varepsilon Nh.$$

(Once more from orthogonality, see above) We'll prove now (0). We may also join $j=0$ whenever $K=W$:

$$\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{m\sim\frac{N}{\ell d}}\sum_{a}W(a\ell)\lll_\varepsilon\sum_{\ell\le 2h}\sum_{(d,q)=1}\sum\frac{1}{q}\left(\frac{N}{\ell d}+1\right)h\lll_\varepsilon Nh$$

and using (2), see Lemma 4, we get (only for $K=W$)

$$T_g(N,h)=2\sum_{\ell\le 2h}\ell\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}\left(1-\cos\frac{2\pi j}{q}[\frac{h}{\ell}]\right)\sum_{0\le|a|\le\frac{h}{\ell}}\left(\left[\frac{h}{\ell}\right]-|a|\right)e_q(ja)+$$

$$+\sum_{\ell\le 2h}\ell\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}B\left(\frac{h}{\ell}\right)\sum_{0\le|a|\le\frac{2h}{\ell}}U_a\left(\frac{h}{\ell}\right)e_q(-ja),$$

(plus negligible remainders), where B and U_a (uniformly on a) are bounded functions. From orthogonality,

$$\frac{1}{q}\sum_{j}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}B\left(\frac{h}{\ell}\right)\sum_{0\le|a|\le\frac{2h}{\ell}}U_a\left(\frac{h}{\ell}\right)e_q(-ja)=B\left(\frac{h}{\ell}\right)\sum_{0\le|a|\le\frac{2h}{\ell}}U_a\left(\frac{h}{\ell}\right)\sum_{\substack{m\sim\frac{N}{\ell d}\\ m\equiv a(q)}}1,$$

whence we get (0), applying orthogonality (and the Lemmas) also on the main term, since for remainders we obtain:

$$\sum_{\ell\le 2h}\ell\sum_{(d,q)=1}\sum g(\ell d)g(\ell q)\frac{1}{q}\sum_{j}\sum_{m\sim\frac{N}{\ell d}}\cos\frac{2\pi jdm}{q}B\left(\frac{h}{\ell}\right)\sum_{0\le|a|\le\frac{2h}{\ell}}U_a\left(\frac{h}{\ell}\right)e_q(-ja)\lll_\varepsilon Nh.$$

We pass to the two bounds for our integrals. Now, from (1) and the well-known formula (Fejér kernel)

$$\sum_{a}S(a)e_q(ja)=\sum_{0\le|a|\le 2h}(2h-|a|)e_q(ja)=\frac{\sin^2\frac{2\pi jh}{q}}{\sin^2\frac{\pi j}{q}},\ \text{ which gives }\ \sum_{a}S(a\ell)e_q(ja)\ge 0\qquad\forall j\ne 0$$

(like in Lemma 4 proof), we have $\forall j \neq 0$, say, (for both $K = W, S$)

$$\widehat{K}\left(\frac{j}{q}\right) \stackrel{def}{=} \sum_a K(a\ell) e_q(ja) \geq 0$$

($\widehat{W}(0) \geq 0$ and $\widehat{S}(0) = \frac{4h^2}{\ell} + \mathcal{O}(h)$, trivially); whence (apart from $\lll_\varepsilon R_g(N,h)$), writing "$*$" for $(d,q)=1$,

$$T_g(N,h) = \sum_{\ell \leq 2h} \sum_{(d,q)=1} \sum g(\ell d) g(\ell q) \frac{1}{q} \sum_{j \neq 0} \sum_{m \sim \frac{N}{\ell d}} \cos \frac{2\pi j d m}{q} \widehat{K}\left(\frac{j}{q}\right) \lll_\varepsilon \max_{D \leq Q} \sum_{\ell \leq 2h} \sum_{q \sim \frac{D}{\ell}} \frac{1}{q} \sum_{d \leq 2q}{}^{*} \sum_{j \neq 0} \frac{1}{\left\| \frac{jd}{q} \right\|} \widehat{K}\left(\frac{j}{q}\right)$$

due to [D, ch. 25]

$$\sum_{m \sim \frac{N}{\ell d}} e_q(jdm) \ll \frac{1}{\left\| \frac{jd}{q} \right\|},$$

having used a dissection argument (over both d, q). Changing variables (with $\overline{d}d \equiv 1(q)$, here) into

$$\sum_{j \neq 0} \frac{1}{\left\| \frac{jd}{q} \right\|} \widehat{K}\left(\frac{j}{q}\right) = \sum_{0 < |j| \leq \frac{q}{2}} \frac{1}{\left\| \frac{j}{q} \right\|} \widehat{K}\left(\frac{j\overline{d}}{q}\right) \xlongequal{K \text{ EVEN}} 2q \sum_{j \leq q/2} \frac{1}{j} \widehat{K}\left(\frac{j\overline{d}}{q}\right)$$

gives

$$T_g(N,h) \lll_\varepsilon \max_{D \leq Q} \sum_{\ell \leq 2h} \sum_{q \sim \frac{D}{\ell}} \sum_{j \leq q/2} \frac{1}{j} \sum_{n \leq 2q}{}^{*} \widehat{K}\left(\frac{jn}{q}\right)$$

which, since

$$\sum_{n \leq 2q}{}^{*} \widehat{K}\left(\frac{jn}{q}\right) \ll \sum_{n \leq 2q} \widehat{K}\left(\frac{jn}{q}\right) = \sum_a K(a\ell) \sum_{n \leq 2q} e_q(jan) = 2q \sum_{\substack{a \\ ja \equiv 0(q)}} K(a\ell)$$

and "flipping" the divisors (say, F, i.e., change t into q/t) in the following

$$2q \sum_{j \leq q/2} \frac{1}{j} \sum_{\substack{a \\ ja \equiv 0(q)}} K(a\ell) = 2 \sum_{\substack{t|q \\ t<q}} q \sum_{\substack{j \leq q/2 \\ (j,q)=t}} \frac{1}{j} \sum_{a \equiv 0(q/t)} K(a\ell) \xlongequal{\text{F}} 2 \sum_{\substack{t|q \\ t>1}} t \sum_{\substack{j \leq t/2 \\ (j,t)=1}} \frac{1}{j} \sum_{a \equiv 0(t)} K(a\ell) =$$

$$= 4h \sum_{\substack{t|q \\ t>1}} t \sum_{\substack{j \leq t/2 \\ (j,t)=1}} \frac{1}{j} + 2 \sum_{\substack{t|q \\ 1<t \leq 2h}} t \sum_{\substack{j \leq t/2 \\ (j,t)=1}} \frac{1}{j} \sum_{\substack{a \neq 0 \\ a \equiv 0(t)}} K(a\ell) \lll_\varepsilon qh + h^2,$$

finally entails

$$T_g(N,h) \lll_\varepsilon Nh + \max_{D \leq Q} \sum_{\ell \leq 2h} \sum_{q \sim \frac{D}{\ell}} (qh + h^2) \lll_\varepsilon Nh + Q^2 h + Qh^2 \Rightarrow (*). \quad \square$$

References

[C-S] Coppola, G. and Salerno, S. - *On the symmetry of the divisor function in almost all short intervals* - Acta Arith. **113** (2004), **no.2**, 189–201. MR 2005a:11144

[D] Davenport, H. - *Multiplicative Number Theory* - Third Edition, GTM 74, Springer, New York, 2000. MR 2001f:11001

[L] Linnik, Ju.V. - *The Dispersion Method in Binary Additive Problems* - Translated by S. Schuur - American Mathematical Society, Providence, R.I. 1963. MR 29#5804

[T] Tenenbaum, G. - *Introduction to Analytic and Probabilistic Number Theory* - Cambridge Studies in Advanced Mathematics, **46**, Cambridge University Press, 1995. MR 97e:11005b

[V] Vinogradov, I.M. - *The Method of Trigonometrical Sums in the Theory of Numbers* - Interscience Publishers LTD, London, 1954. MR 15, 941b

Dr.Giovanni Coppola
DIIMA - Università degli Studi di Salerno
84084 Fisciano (SA) - ITALY
e-mail : gcoppola@diima.unisa.it